2022/10/06

藪友良

微分について

ここでは、微分と偏微分について説明します。微分と偏微分について、十分に理解している方は、この資料を読む必要はありません。

B.1 微分

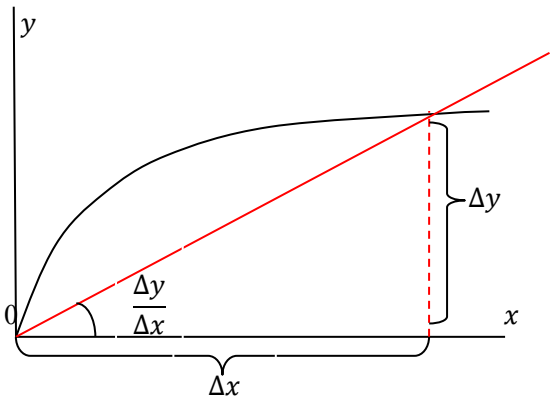
ある値xに対して、別の値$f(x)$を対応させる関係を 1 変数の関数といいます。たとえば、$f(x)=4+2x$、$f(x)=3x^2$といった 2 つの式は、それぞれ変数xを$f(x)$に対応させる関係であり、1 変数の関数の例です。

微分とは、「関数$f(x)$の接線の傾き」であり、「xを微小に変化させたとき、$f(x)$がどれだけ変化するか」を表します。微分は、$f'(x)$や$df(x)/dx$と表記されます。

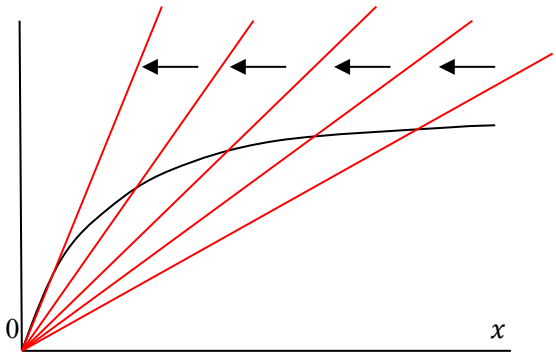
ここで、$y=f(x)$としましょう。図 B-1(a)では、$y=f(x)$を示しています。図をみると、xが増えるとyも増加しますが、その増加幅は少しずつ減少する関数であることが分かります。ここで、原点 0 から、xはΔxだけ変化すると、yはΔyだけ変化します(Δは変化量を示す記号です。すなわち、Δxはxの変化量、Δyはyの変化量を表します)。そして、図 B-1(a)において、Δxを底辺、Δyを高さとした三角形の傾きは、$\Delta y/\Delta x$となります。ここで、図 B-1(b)のように、底辺Δxを小さくしていくと、三角形の傾き$\Delta y/\Delta x$は、原点 0 で評価した関数の接線の傾きに近づいていきます。そして、Δxが 0 に非常に近い値であるとき、$\Delta y/\Delta x$は原点 0 で評価した関数の接線の傾きになります。これが$x=0$で評価した微分となります。

図 B-1： 原点 0 での微分のイメージ

(a) $\Delta y/\Delta x$



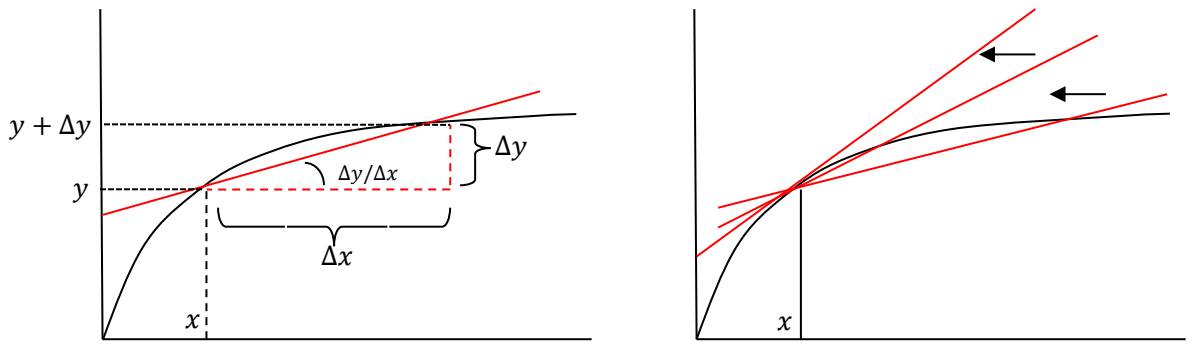
(b) Δxを小さくする



関数の接線の傾きは、xをどの値で評価するかによって異なります。たとえば、xは 0 ではないとします。図 B-2(a)から、ある値$x(\neq 0)$から$x+\Delta x$に変化したとき、yは$y+\Delta y$に変化します。ここで、$\Delta y/\Delta x$は、Δxが大きいため、関数の傾きではありません。しかし、図 B-2(b)のように、Δxを 0 に近づけていくと、$\Delta y/\Delta x$は、xで評価した関数の接線の傾きに近づいていきます。これがxで評価した微分となります。

図 B-2　一般的な微分のイメージ

(a)　$\Delta y/\Delta x$　　**(b)　Δxを小さくする**



微分のイメージが理解できたでしょうか。繰り返しになりますが、微分とは、関数$f(x)$の「接線の傾き」であり、「xが変化したとき、$y=f(x)$がどれぐらい変化するか」を表します。数式を用いると、微分は

$$\frac{dy}{dx} = lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x}$$

$$= lim_{\Delta x \to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

と定義されます。ここで、$lim_{\Delta x \to 0}$とは、xの変化が非常に小さい(Δxが 0 に近い)状態を表します。

B.2　微分の公式

ここで、$y=x^2$という関数を考えましょう。このとき、微分は$2x$です。

$$y = x^2 \qquad \Rightarrow \qquad \frac{dy}{dx} = 2x$$

[証明]　xから$x+\Delta x$に微小に変化したとします。このとき、$\Delta y/\Delta x$は、

$$\frac{f(x+\Delta x)-f(x)}{\Delta x} = \frac{(x+\Delta x)^2 - x^2}{\Delta x}$$

$$= \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} = 2x + \Delta x$$

となります。ここでΔxを 0 に近づけると、第 2 項目が消えて$2x$となります。[終]

一般的には、$y = x^n$という関係を考えると、微分は次のとおりです。

$$y = x^n \qquad \Rightarrow \qquad \frac{dy}{dx} = nx^{n-1}$$

[証明]　ここで、xから$x_* = x + \Delta x$に変化した状況を考えます(つまり、$\Delta x = x_* - x$)。よって、微分の定義から、次のようになります。

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{x_*^n - x^n}{x_* - x}$$

ここで、右辺の別表現を考えます。まず、$n = 2$ならば、

$$\frac{x_*^2 - x^2}{x_* - x} = x_* + x$$

となります。同様に、$n = 3$、$n = 4$の場合を考えよう。このとき、

$$\frac{x_*^3 - x^3}{x_* - x} = x_*^2 + x_* x + x^2$$

$$\frac{x_*^4 - x^4}{x_* - x} = x_*^3 + x x_*^2 + x^2 x_* + x^3$$

となります。両辺に$x_* - x$を掛けて、両辺が等式となることを確認してください。これを一般化すると、任意のnに対して、

$$\frac{x_*^n - x^n}{x_* - x} = x_*^{n-1} + x x_*^{n-2} + x^2 x_*^{n-3} + \cdots + x^{n-2} x_* + x^{n-1}$$

と表現できます。ただし、上式では、x_*^{n-j}の乗数$n - j$が負なら、その項はすべて無視するとします。この関係を用いると、

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{x_*^n - x^n}{x_* - x} = x_*^{n-1} + x x_*^{n-2} + x^2 x_*^{n-3} + \cdots + x^{n-2} x_* + x^{n-1}$$

となります。ここで、Δxを 0 に近づけると、$x_* = x + \Delta x$は、xに限りなく近づいていきます。よって、微分は、上式のx_*をxで置き換えた式、つまり、

$$x^{n-1} + x x^{n-2} + x^2 x^{n-3} + \cdots + x^{n-1} = x^{n-1} + x^{n-1} + x^{n-1} + \cdots + x^{n-1} = nx^{n-1}$$

となります。　[終]

ここでは、さらに上の公式を一般化します。任意の定数aとbを考えて、関数$y = a + bx^n$を定義します。この関数を微分すると、

$$y = a + bx^n \quad \Rightarrow \quad \frac{dy}{dx} = bnx^{n-1}$$

となります。任意の定数であるaは、xに依存してないため、微分の式に表れていません。

[証明] xから$x_* = x + \Delta x$に変化した状況を考えます(つまり、$\Delta x = x_* - x$)。よって、微分の定義から、次のようになります。

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{(a + bx_*^n) - (a + bx^n)}{x_* - x} = b\left(\frac{x_*^n - x^n}{x_* - x}\right)$$

ここで、$\Delta x = x_* - x$を 0 に近づけると、$(x_*^n - x^n)/(x_* - x)$はnx^{n-1}になるため、微分はbnx^{n-1}です。 [終]

例 B-1：微分の数値例

$y = a$とします。このとき、aはxに依存していないため、

$$\frac{dy}{dx} = 0$$

となります。これは公式で$b = 0$とした場合に該当します。つまり、定数の微分は 0 です。

次に、$y = bx$とします。このとき、公式より、

$$\frac{dy}{dx} = bx^0 = b$$

となります。つまり、xで微分すると定数bとなります。

最後に、$y = 3 + 4x^3$とします。このとき、公式より、次のようになります。

$$\frac{dy}{dx} = 4 \times 3x^2 = 12x^2$$

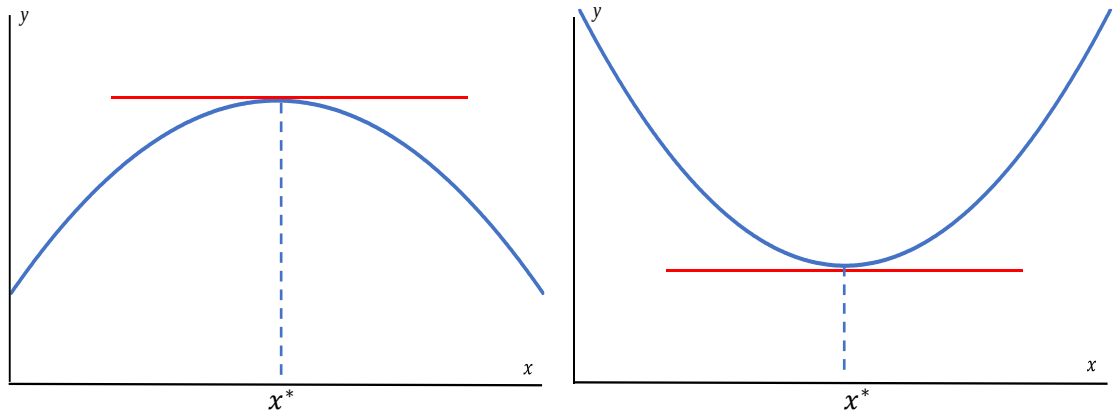
B.3　関数の最小化と最大化

微分を用いることで何が分かるでしょうか。微分は関数の傾き、つまり、「xが変化したときのyの変化」を教えてくれるので、それ自体が有用な情報を与えてくれます。それ以外にも、微分を用いることで、関数$f(x)$の最大値や最小値を見つけることもできます。

図 B-3 では、2 つの関数$f(x)$を描いています。これをみると、関数の最大値と最小値において、関数の傾きが 0 になっているのがみてとれます。実は、関数$f(x)$の微分が 0 となるポイントx^*とは、関数を最大化もしくは最小化するポイントx^*にもなっているのです[1]。

図 B-3　微分が 0 となるポイントx^*

(a) 最大値　　**(b)　最小値**



ここで$y = x^2$としましょう。公式より$dy/dx = 2x$です。図 B-4(a)には、$y = x^2$という関数を、図 B-4(b)には、微分(接線の傾き)である$2x$を描きました。当然ですが、$2x$という関数は、xが負のとき負である一方、xが正のとき正です。これは$x < 0$で微分が負に、$x > 0$で微分が正となることを意味します。ここで、$x = 0$で微分は 0 となるため、この関数は、$x = 0$で最小値となります(厳密には、$2x = 0$を満たすxは、$x = 0$と求められる)。

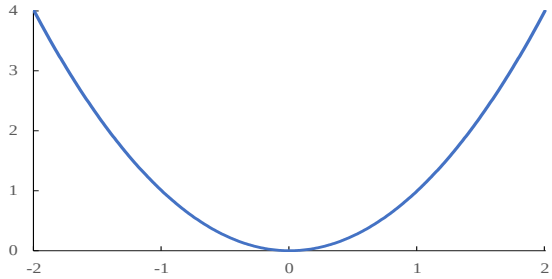
これまでの議論から、関数を微分して 0 と置いた式をxについて解くことで、関数を最大化もしくは最小化するポイントx^*を求められることが理解できました。では、こうして求めた点x^*が、最大化と最小化のどちらであるのかを、ど

[1] これは局所的に最大化もしくは最小化されている可能性もあります。

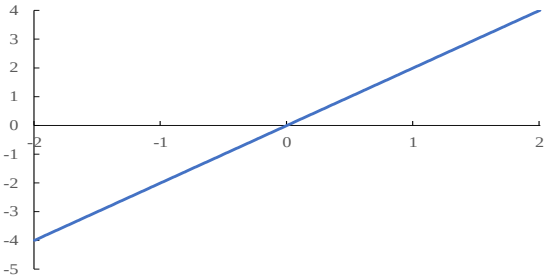
うしたら判断できるのでしょうか。もちろん関数を図示すれば分かりますが、図に描くのも面倒です。別の方法として、微分して 0 と置いた点x^*から微小にxを変化させ、関数の値が減少すれば、x^*は最大化点と判断することができます。

図 B-4　関数$y = x^2$

(a)　$y = x^2$



(b)　微分$2x$



B.4　合成関数

ここで、$y = f(x)$とし、またxは別の変数zの関数とします(これを$g(z)$と表す)。このとき、yはzの関数となっており、$f(g(z))$を$f(x)$と$g(z)$の合成関数と呼びます。たとえば、$y = (3 + 2z)^2$とすると、これは$y = x^2$、$x = 3 + 2z$という合成関数と考えられます(つまり、$y = x^2$は、$f(x) = x^2$と$g(z) = 3 + 2z$の合成関数)。

B.4.1　合成関数の微分

合成関数の微分は、関数$f(g(z))$の傾きとなります。これは合成関数の微分の公式によって、

$$\frac{dy}{dz} = \frac{dy}{dx}\frac{dx}{dz}$$

と求めることができます。この公式は、言葉で説明すると、

$$\begin{pmatrix} z\text{が変化したときの} \\ y\text{の変化量} \end{pmatrix} = \begin{pmatrix} x\text{が変化したときの} \\ y\text{の変化量} \end{pmatrix} \times \begin{pmatrix} z\text{が変化したときの} \\ x\text{の変化量} \end{pmatrix}$$

となる、といえます。これは直観的に理解できる式ではないでしょうか。

例 B-2：合成関数の微分の数値例

$y = (3 + 2z)^2$とし、これをzに関して微分したいとします。この関数は、$y = x^2$と$x = 3 + 2z$ との合成関数です。また、

$$\frac{dy}{dx} = 2x、\frac{dx}{dz} = 2$$

となるため、合成関数の微分の公式から、dy/dzは次のとおりです。

$$\frac{dy}{dz}=\frac{dy}{dx}\frac{dx}{dz}=4x=4\times(3+2z)=12+8z$$

B.5　偏微分

これまで、ある変数xが、ある値$f(x)$に対応する 1 変数の関数である場合について考えてきました。これに対して、もし 2 変数x_1とx_2が、ある値$f(x_1,x_2)$に対応するならば、これは 2 変数の関数といいます。同様にして、2 つ以上の変数の関数も定義できます。

ここで、x_1に関する偏微分とは、「x_2は固定した値として、x_1に関して微分をとったもの」です。これは、

$$\frac{\partial y}{\partial x_1}=lim_{\Delta x_1\to 0}\frac{\Delta y}{\Delta x_1}$$

$$=lim_{\Delta x_1\to 0}\frac{f(x_1+\Delta x_1,x_2)-f(x_1,x_2)}{\Delta x_1}$$

となります。ここで、x_2は固定した値とし、x_1だけが$x_1+\Delta x_1$に変化しています。

偏微分では、微分と明確に区別するため、記号として、dではなく∂(ラウンド・デルタと呼ぶ)を用いています。偏微分記号を用いるときは、変数が複数あり、そのうち 1 つを微小に変化させる状況を考えています。偏微分と聞くと難しそうですが、他の変数を固定した値とみなして、微分をとるだけなので、とても簡単です。

例 B-3：偏微分の数値例

ここで、$y=x_1^2x_2$を、x_1で偏微分します。x_2を固定した値とみなすので、偏微分は以下となります。

$$\frac{\partial y}{\partial x_1}=2x_1x_2$$

次に、$y=x_1^3+2x_1x_2-4x_2$を、x_1で偏微分します。x_2を固定した値とみなすので、偏微分は、次の通りです。

$$\frac{\partial y}{\partial x_1}=3x_1^2+2x_2$$